

# An Approximate Projected Consensus Algorithm for Computing Intersection of Convex Sets\*

Youcheng Lou<sup>†</sup>, Guodong Shi<sup>‡</sup>, Karl Henrik Johansson and Yiguang Hong

## Abstract

In this paper, we propose an approximate projected consensus algorithm for a network to cooperatively compute the intersection of convex sets. Instead of assuming the exact convex projection proposed in the literature, we allow each node to compute an approximate projection and communicate it to its neighbors. The communication graph is directed and time-varying. Nodes update their states by weighted averaging. Projection accuracy conditions are presented for the considered algorithm. They indicate how much projection accuracy is required to ensure global consensus to a point in the intersection set when the communication graph is uniformly jointly strongly connected. We show that  $\pi/4$  is a critical angle error of the projection approximation to ensure a bounded state. A numerical example indicates that this approximate projected consensus algorithm may achieve better performance than the exact projected consensus algorithm in some cases.

**Keywords:** Multi-agent systems, approximate projection, intersection computation, optimal consensus

## 1 Introduction

In recent years, dynamics and control on large-scale networks have drawn increasing research attention in different areas including engineering, computer science, and social science. Cooperative control of a group of autonomous agents fully employs local information exchange and distributed protocol design to accomplish collective tasks such as consensus, formation, and aggregation [7, 8, 18, 15, 16, 34, 19, 17, 11, 12]. Moreover, in parallel computation, load-balance problems require realtime balance of the load from different computing resources [9, 10]. Additionally, a central problem of opinion dynamics in social networks is how the agreement is achieved via individual belief exchange processes [13, 14]. A fundamental question in these problems is, how consensus can be guaranteed based on local information exchange, time-varying node interconnection and limited knowledge of the global objective.

---

\*This work is partially supported by the NNSF of China under Grant 61174071, the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

<sup>†</sup>Y. Lou and Y. Hong are with Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: [louyoucheng@amss.ac.cn](mailto:louyoucheng@amss.ac.cn), [yghong@iss.ac.cn](mailto:yghong@iss.ac.cn)

<sup>‡</sup>G. Shi and K. H. Johansson are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden. Email: [guodongs@kth.se](mailto:guodongs@kth.se), [kallej@ee.kth.se](mailto:kallej@ee.kth.se)

Various distributed optimization problems arise for consensus with particular optimization purpose in practice. Minimizing a sum of convex functions, where each component is known only to a particular node, has attracted much attention, due to its simple formulation and wide applications [22, 20, 21, 26, 27, 23, 32, 31, 29, 30, 33, 25, 24]. The key idea is that properly designed distributed control protocols or computation algorithms can lead to a collective optimization, based on simple exchanged information and individual optimum observation. Subgradient-based incremental methods were established via deterministic or randomized iteration, where each node is assumed to be able to compute a local subgradient value of its objective function [20, 21, 26, 22, 25, 24]. Non-subgradient-based methods also showed up in the literature. For instance, a non-gradient-based algorithm was proposed, where each node starts at its own optimal solution and updates using a pairwise equalizing protocol [29, 30], and later an augmented Lagrangian method was introduced in [33].

In particular, if the optimal solution set of its own objective can be obtained for each node, the considered optimization problem is then converted to a set intersection computation problem when we additionally assume there is a nonempty intersection among all solution sets [32, 31, 27]. In fact, convex intersection computation problem is a classical problem in the optimization study [35, 36, 37]. The so-called “alternating projection algorithm” was a standard centralized solution, where projection is carried out alternatively onto each set [35, 36, 37]. Then the “projected consensus algorithm” was presented as a decentralized version of alternating projection algorithm, where each node alternatively projects onto its own set and averages with its neighbors, and comprehensive convergence analysis was given for this projected algorithm under time-varying directed interconnections in [27]. Following this work, a flip-coin algorithm was introduced when each node randomly chooses projection or averaging by Bernoulli processes, and almost sure convergence was shown for the system to reach an optimal consensus in [32]. A dynamical system solution was given in [31], where the network reaches a global optimal consensus by a simple continuous-time control. In all these algorithms, each node needs to know the exact convex projection of its current state onto its objective set [32, 31, 27].

However, in practice, the exact convex projection is usually hard to compute due to the common environmental noise and computation inaccuracy. In this paper, we therefore propose an approximate projected consensus algorithm to solve the convex intersection computation problem. Instead of assuming the exact convex projection, we allow each node to just compute an approximate projection point which locates in the intersection of the convex cone generated by the current state and all directions with the exact projection direction less than some angle and the half-space containing the current state with its boundary being a supporting hyperplane to its own set at its exact projection point onto its set. The communication graph is supposed to be directed and time-varying. With uniformly jointly strongly connected conditions, we show that the whole network can achieve a global consensus within the intersection of all convex sets when sufficient projection accuracy can be guaranteed. For a special approximate projection case when the nodes can get the exact direction of the projection, a necessary and sufficient condition is given on how much projection accuracy is critical to ensure a global intersection computation. A numerical example is also given, and surprisingly, the approximate projected consensus algorithm sometimes achieves better performance for convergence than the exact projected consensus algorithm.

The paper is organized as follows. Section 2 gives some basic concepts on graph theory and convex analysis. Section 3 introduces the network model and formulates the problem of interest. Section 4 presents the main results and convergence analysis for the considered algorithm. Section 5 focuses on the discussion on the critical angle error of the approximate projection. Section 6 gives a numerical example and finally, Section 7 shows some concluding remarks.

## 2 Preliminaries

In this section, we introduce preliminary knowledge on graph theory [5] and convex analysis [1].

A directed graph (digraph)  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$  consists of node set  $\mathcal{V} = \{1, 2, \dots, n\}$ , arc set  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  and an adjacency matrix  $A = [a_{ij}]_{n \times n}$  with nonnegative adjacency elements  $a_{ij}$ . The element  $a_{ij}$  of matrix  $A$  associated with arc  $(i, j)$  is positive if and only if  $(i, j) \in \mathcal{E}$ .  $\mathcal{N}_i$  denotes the set of neighbors of node  $i$ , that is,  $\mathcal{N}_i = \{j \in \mathcal{V} | (i, j) \in \mathcal{E}\}$ . In this paper, we assume  $(i, i) \in \mathcal{E}$  for all  $i$ . A path from  $i$  to  $j$  in digraph  $\mathcal{G}$  is a sequence  $(i_0, i_1), (i_1, i_2), \dots, (i_{p-1}, i_p)$  of arcs with  $i_0 = i$  and  $i_p = j$ .  $\mathcal{G}$  is said to be strongly connected if there exists a path from  $i$  to  $j$  for each pair of nodes  $i, j \in \mathcal{V}$ .

A function  $f(\cdot) : R^m \rightarrow R$  is said to be convex if  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$  for any  $x, y \in R^m$  and  $0 < \lambda < 1$ . A function  $f$  is said to be concave if  $-f$  is convex. A set  $K \subseteq R^m$  is said to be convex if  $\lambda x + (1 - \lambda)y \in K$  for any  $x, y \in K$  and  $0 < \lambda < 1$  and is said to be a convex cone if  $\lambda_1 x + \lambda_2 y \in K$  for any  $x, y \in K$  and  $\lambda_1, \lambda_2 \geq 0$ . For a closed convex set  $K$  in  $R^m$ , we can associate to any  $x \in R^m$  a unique element  $P_K(x) \in K$  satisfying  $|x - P_K(x)| = \inf_{y \in K} |x - y|$ , which is denoted as  $|x|_K$ , where  $|\cdot|$  denotes the Euclidean norm and  $P_K$  is the projection operator onto  $K$ .

For a closed convex set  $K$ , if  $x \notin K$ , then by the supporting hyperplane theorem, there is a supporting hyperplane to  $K$  at  $P_K(x)$ . The angle between vectors  $a$  and  $b$  is denoted as  $\text{Ang}(a, b) \in [0, \pi]$  for which  $\cos \text{Ang}(a, b) = \langle a, b \rangle / (|a||b|)$ , where  $\langle a, b \rangle$  denotes the Euclidean inner product of vectors  $a$  and  $b$ .

We cite a lemma for the following analysis (see Example 3.16 in [3], pp. 88).

**Lemma 2.1**  $f(z) = |z|_K$  is a convex function, where  $K$  is a closed convex set in  $R^m$ .

The following properties hold for the projection operator  $P_K$ . Here (i) is the standard non-expansiveness property for convex projection; (ii) comes from exercise 1.2 (c) in [2] (pp. 23) and (iii) is a special case of proposition 1.3 in [2] (pp. 24).

**Lemma 2.2** Let  $K$  be a closed convex set in  $R^m$ . Then

- (i)  $|P_K(x) - P_K(y)| \leq |x - y|$  for any  $x$  and  $y$ ;
- (ii)  $||x|_K - |y|_K| \leq |x - y|$  for any  $x$  and  $y$ ;
- (iii)  $P_K(\lambda x + (1 - \lambda)P_K(x)) = P_K(x)$  for any  $x$  and  $0 < \lambda < 1$ .

The next lemma can be found in [32].

**Lemma 2.3** Let  $K$  and  $K_0 \subseteq K$  be two closed convex sets. We have

$$|P_K(x)|_{K_0}^2 + |x|_K^2 \leq |x|_{K_0}^2 \text{ for any } x.$$

### 3 Problem Formulation

In this section, we introduce the intersection computation problem and the approximate projected consensus algorithm.

Consider a network consisting of  $n$  agents with node set  $\mathcal{V} = \{1, 2, \dots, n\}$ . Each node  $i$  is associated with a set  $X_i \subseteq R^m$  and set  $X_i$  is known only by node  $i$ . The intersection of all these sets is nonempty, i.e.,  $\bigcap_{i=1}^n X_i \neq \emptyset$ . Let us denote  $X_0 = \bigcap_{i=1}^n X_i$ . The target of the network is to find a point in  $X_0$  in a distributed way. For  $X_i, i = 1, \dots, n$ , we use the following assumption:

**A1** (*Convexity*)  $X_i, i = 1, \dots, n$ , are closed convex sets.

**Remark 3.1** *The intersection computation problem can be equivalently converted into the following distributed optimization problem: the objective of this group of  $n$  agents is not only to achieve a consensus, but also to cooperatively solve*

$$\min_{x \in R^m} \sum_{i=1}^n f_i(x),$$

where  $f_i : R^m \rightarrow R$  is the convex cost function of agent  $i$  and can be known only by agent  $i$ . Here  $X_i = \{y | f_i(y) = \min_{x \in R^m} f_i(x)\}, 1 \leq i \leq n$  are closed convex sets, which are assumed to be nonempty and have a nonempty intersection.

#### 3.1 Communication Graphs

The communication over the network is modeled as a sequence of directed graphs,  $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}(k), A(k)), k \geq 0$ . We say node  $j$  is a neighbor of node  $i$  at time  $k$  if there is an arc  $(i, j) \in \mathcal{E}(k)$ , where  $a_{ij}(k)$  represents its weight. Let  $\mathcal{N}_i(k)$  denote the set of neighbors of agent  $i$  at time  $k$ . We introduce an assumption on the weights [26, 32].

**A2** (*Weights Rule*) (i)  $\sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) = 1$  for all  $i$  and  $k$ .

(ii) There exists a constant  $0 < \eta < 1$  such that  $a_{ij}(k) \geq \eta$  for all  $i, k$  and  $j \in \mathcal{N}_i(k)$ .

For the connectivity of the communication graphs, we introduce the following assumption [31, 27].

**A3** (*Connectivity*) The communication graph is uniformly jointly strongly connected (*UJSC*), i.e., there exists a positive integer  $T$  such that  $\mathcal{G}([k, k+T))$  is strongly connected for  $k \geq 0$ , where  $\mathcal{G}([k, k+T))$  denotes the union graph with node set  $\mathcal{V}$  and arc set  $\bigcup_{k \leq s < k+T} \mathcal{E}(s)$ .

#### 3.2 Approximate Projection

Projection methods have been widely used to solve various problems, including projected consensus [27], the convex intersection computation [36, 37] and distributed computation [4]. In the most literature, the projection point  $P_K(z)$  of  $z$  onto closed convex set  $K$  is required to achieve desired convergence, but in practice it is hard to be obtained and often is computed approximately. Here is the definition of approximate projection.

**Definition 3.1** Suppose  $K \subseteq R^m$  is a closed convex set and  $0 < \theta < \pi/2$ . Define

$$\begin{aligned}\mathbf{C}_K(v, \theta) &= v + \{z \mid \langle z, P_K(v) - v \rangle \geq |z||v|_K \cos \theta\}; \\ \mathbf{H}_K^+(v) &= \{z \mid \langle v - P_K(v), z \rangle \geq \langle v - P_K(v), P_K(v) \rangle\}.\end{aligned}$$

as a convex cone and a half-space, respectively. The approximate projection  $\mathcal{P}_K^a(v, \theta)$  of point  $v$  onto  $K$  with approximate angle  $\theta$  is defined as the following set:

$$\mathcal{P}_K^a(v, \theta) = \begin{cases} \mathbf{C}_K(v, \theta) \cap \mathbf{H}_K^+(v), & \text{if } v \notin K \\ \{v\}, & \text{if } v \in K \end{cases} \quad (1)$$

In fact,  $\mathbf{C}_K(v, \theta) - v$  is a convex cone generated by all vectors having angle with  $P_K(v) - v$  less than  $\theta$  and  $\mathbf{H}_K^+(v)$  is the half-space containing point  $v$  with

$$\mathbf{H}_K(v) := \{z \mid \langle v - P_K(v), z \rangle = \langle v - P_K(v), P_K(v) \rangle\}$$

being a supporting hyperplane to  $K$  at  $P_K(v)$ .

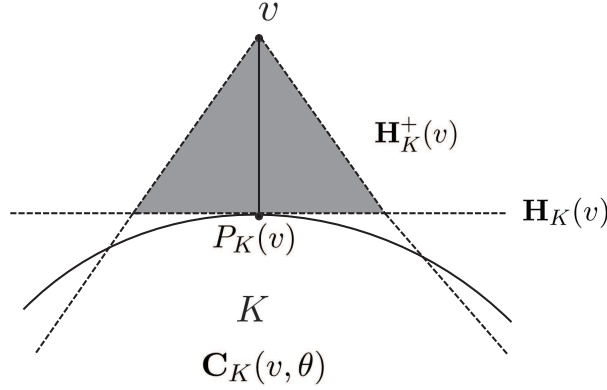


Figure 1: The set marked by the shaded area is the approximate projection of  $v$  onto  $K$ .

**Definition 3.2** The supporting approximate projection  $\mathcal{P}_K^{sa}(v, \theta)$  of point  $v$  onto  $K$  with approximate angle  $\theta$  is defined as

$$\mathcal{P}_K^{sa}(v, \theta) = \begin{cases} \mathbf{C}_K(v, \theta) \cap \mathbf{H}_K(v), & \text{if } v \notin K \\ \{v\}, & \text{if } v \in K \end{cases}$$

**Remark 3.1** Exact projection may be difficult to obtain in practice, due to possible data quantization and limitation of the estimate. This is why we introduce approximate projection, as a relaxed estimate to the real projection value.

According to Definition 3.2, for any  $y \in \mathcal{P}_K^a(v, \theta)$ , we can associate  $y$  with  $\hat{y} \in \mathcal{P}_K^{sa}(v, \theta)$  such that

$$y = (1 - \beta)v + \beta\hat{y} \text{ for some } 0 \leq \beta \leq 1. \quad (2)$$

Moreover, it is easy to see that if  $y \neq v$ ,  $\hat{y}$  satisfying (2) is unique.

### 3.3 Distributed Iterative Algorithm

To solve the intersection computation problem, we propose the following approximate projected consensus algorithm:

$$x_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) P_j^a(k), \quad i = 1, \dots, n, \quad (3)$$

where  $P_i^a(k) \in \mathcal{P}_{X_i}^a(x_i(k), \theta_k)$  for all  $i$  and  $k$ .

According to the definition of supporting approximate projection, there exist  $0 \leq \alpha_{i,k} \leq 1$  and  $P_i^{sa}(k) \in \mathcal{P}_{X_i}^{sa}(x_i(k), \theta_k)$  such that

$$P_i^a(k) = (1 - \alpha_{i,k})x_i(k) + \alpha_{i,k}P_i^{sa}(k). \quad (4)$$

Combining with (3) and (4), we have

$$x_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \left( (1 - \alpha_{j,k})x_j(k) + \alpha_{j,k}P_j^{sa}(k) \right), \quad (5)$$

where if  $x_i(k) \notin X_i$ ,  $P_i^{sa}(k) \in \mathbf{H}_{X_i}(x_i(k))$  and  $\text{Ang}(P_i^{sa}(k) - x_i(k), P_{X_i}(x_i(k)) - x_i(k)) \leq \theta_k$ .

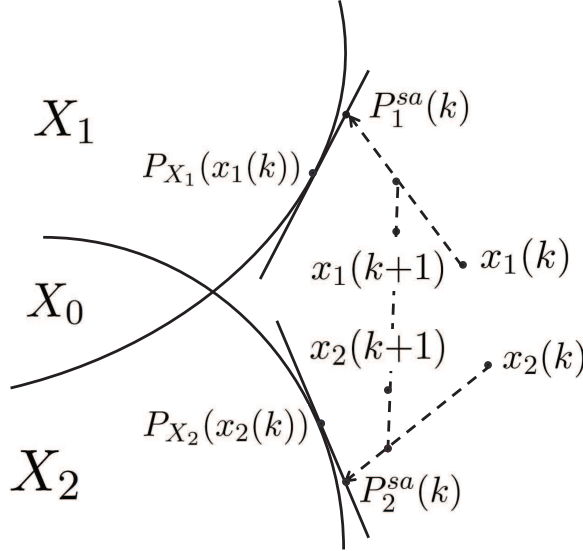


Figure 2: The approximate projected consensus algorithm.

**Remark 3.2** Note that the “projected consensus algorithm” presented in [27] is a special case with  $\alpha_{i,k} \equiv 1$  and  $\theta_k \equiv 0$  of the approximate projected consensus algorithm discussed in this paper.

We illustrate the iteration process of the algorithm in Figure 2. For the approximate angle  $\theta_k$ , we use the following assumption.

**A4**  $0 \leq \theta_k \leq \theta^* < \pi/2$  for all  $k$ .

In this paper, we are interested in whether an optimal consensus can be achieved or not, as defined in the following definition.

**Definition 3.3** A global optimal consensus is achieved for the approximate projected consensus algorithm if, for any initial condition  $x(0) \in R^{nm}$ , there exists  $x^* \in X_0$  such that

$$\lim_{k \rightarrow \infty} x_i(k) = x^*, \quad i = 1, \dots, n.$$

Note that a global optimal consensus  $x^*$  necessarily belongs to the intersection set  $X_0$ . Hence, a possible algorithm to find a point in  $X_0$  is to employ the approximate projected consensus algorithm (3). We next discuss the convergence of this algorithm.

## 4 Main Results and Convergence Analysis

In this section, we present the main results and convergence analysis.

Denote  $\alpha_k^- = \min_{1 \leq i \leq n} \alpha_{i,k}$  and  $\alpha_k^+ = \max_{1 \leq i \leq n} \alpha_{i,k}$ ,  $k \geq 0$ .

**Theorem 4.1** Suppose **A1–A4** hold. Global optimal consensus is achieved for the approximate projected consensus algorithm if  $\sum_{k=0}^{\infty} \alpha_k^- = \infty$  and  $\sum_{k=0}^{\infty} \alpha_k^+ \theta_k < \infty$ .

To investigate the necessity of the divergence condition in Theorem 4.1, we impose another assumption on the boundedness of the  $n$  sets  $X_i, i = 1, \dots, n$ .

**A5** (Bounded Sets)  $X_i, i = 1, \dots, n$ , are bounded sets.

Then the following conclusion holds showing a necessary projection accuracy condition.

**Theorem 4.2** Suppose **A1–A5** hold,  $\theta_k \equiv 0$  and  $\alpha_k^+ < 1$  for all  $k$ . Global optimal consensus is achieved for the approximate projected consensus algorithm only if  $\sum_{k=0}^{\infty} \alpha_k^+ = \infty$ . In fact, if  $\sum_{k=0}^{\infty} \alpha_k^+ < \infty$ , then for initial condition  $x_i(0) = z^*, i = 1, \dots, n$  with  $|z^*|_{X_0} > \sum_{k=0}^{\infty} \alpha_k^+ d^* / \prod_{k=0}^{\infty} (1 - \alpha_k^+)$ , there is  $y^* = y^*(z^*) \notin X_0$  such that  $\lim_{k \rightarrow \infty} x_i(k) = y^*$  for all  $i$ , where  $d^* = \sup_{\omega_1, \omega_2 \in \bigcup_{i=1}^n X_i} |\omega_1 - \omega_2|$ .

**Remark 4.1** Suppose **A1–A5** hold,  $\theta_k \equiv 0$  and there exists a sequence  $\{\alpha_k\}_{k=0}^{\infty}$  with  $\alpha_k < 1$  for all  $k$  such that  $\alpha_{i,k} = \alpha_{j,k} = \alpha_k$  for all  $i, j$  and  $k$ . Then from Theorems 1 and 2, we have that global optimal consensus is achieved for the approximate projected consensus algorithm if and only if  $\sum_{k=0}^{\infty} \alpha_k = \infty$ .

**Remark 4.2** Compared to the convergence results given in [27], Theorems 4.1 and 4.2 do not require the doubly stochastic assumption on the weights  $a_{ij}(k)$  ( $\sum_{j=1}^n a_{ij}(k) = \sum_{j=1}^n a_{ji}(k) = 1$  for all  $i, k$ ). This is important because double stochasticity is hard to guarantee for the arc weights in a distributed way, especially when the communication is directed.

Moreover, the connectivity assumption in [27] requires that  $\mathcal{G}([k, k+T))$  is a fixed graph for sufficiently large  $k$ , which is more restrictive than our UJSC assumption. However, the assumption in [27] can be relaxed to UJSC graphs, as indicated in a comment by the authors.

### 4.1 Lemmas

We establish several useful lemmas in this subsection.

Let  $\{x_i(k)\}_{k=0}^\infty$  be the states of node  $i$  generated by (3),  $i = 1, \dots, n$ . Denote  $|x(k)|_{X_0} = (|x_1(k)|_{X_0} \dots |x_n(k)|_{X_0})^T$ ,  $y(k) = (y_1(k) \dots y_n(k))^T$  with  $y_i(k) = |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2}$ . Denoting  $D_k = \text{diag}\{\alpha_{1,k}, \alpha_{2,k}, \dots, \alpha_{n,k}\}$ , the following lemma holds.

**Lemma 4.1** *Suppose A1 holds. For all  $k \geq s$ , we have*

$$|x(k+1)|_{X_0} \leq A(k)|x(k)|_{X_0} - A(k)D_k y(k) + \tan \theta_k A(k)D_k |x(k)|_{X_0}. \quad (6)$$

Proof. According to Lemma 2.1, (5) implies

$$|x_i(k+1)|_{X_0} \leq \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \left( (1 - \alpha_{i,k})|x_j(k)|_{X_0} + \alpha_{i,k}|P_j^{sa}(k)|_{X_0} \right). \quad (7)$$

By Lemma 2.2 (ii), we have

$$|P_j^{sa}(k)|_{X_0} \leq |P_j^{sa}(k) - P_{X_j}(x_j(k))| + |P_{X_j}(x_j(k))|_{X_0}. \quad (8)$$

The definition of  $P_j^{sa}(k)$  ensures that

$$|P_j^{sa}(k) - P_{X_j}(x_j(k))| \leq \tan \theta_k |x_j(k)|_{X_j}. \quad (9)$$

Moreover, it follows from Lemma 2.3 that for any  $j \in \mathcal{V}$ ,

$$|P_{X_j}(x_j(k))|_{X_0} \leq \sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2}. \quad (10)$$

It follows from (7), (8), (9), (10) and the relation  $|x_j(k)|_{X_j} \leq |x_j(k)|_{X_0}$  that

$$\begin{aligned} |x_i(k+1)|_{X_0} &\leq \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \left( (1 - \alpha_{j,k})|x_j(k)|_{X_0} + \alpha_{j,k} \sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2} \right) \\ &\quad + \tan \theta_k \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \alpha_{j,k} |x_j(k)|_{X_0} \end{aligned} \quad (11)$$

$$\begin{aligned} &= \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \left( |x_j(k)|_{X_0} - \alpha_{j,k} (|x_j(k)|_{X_0} - \sqrt{|x_j(k)|_{X_0}^2 - |x_j(k)|_{X_j}^2}) \right) \\ &\quad + \tan \theta_k \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \alpha_{j,k} |x_j(k)|_{X_0}. \end{aligned} \quad (12)$$

Then the conclusion follows.  $\square$

It is easy to find that  $\tan \theta \leq (\tan \theta^* / \theta^*) \theta$  for  $0 \leq \theta \leq \theta^*$ . Thus, if  $\sum_{k=0}^\infty \alpha_k^+ \theta_k < \infty$ , then  $\sum_{k=0}^\infty \alpha_k^+ \tan \theta_k < \infty$ .

**Lemma 4.2** *Suppose A1 and A4 hold. If  $\sum_{k=0}^\infty \alpha_k^+ \theta_k < \infty$ ,  $\{x_i(k)\}_{k=0}^\infty$  is bounded for all  $i$ .*

Proof. Lemma 1 (b) in [27] states that for closed convex  $K$  and any  $x \in R^m$ ,

$$|P_K(x) - y|^2 \leq |x - y|^2 - |x|_K^2 \text{ for all } y \in K. \quad (13)$$



By taking  $K = X_j$  and  $z \in X_0 \subseteq X_j$ , (13) leads to

$$|P_{X_j}(x_j(k)) - z| \leq \sqrt{|x_j(k) - z|^2 - |x_j(k)|_{X_j}^2}. \quad (14)$$

By considering  $|x_i(k+1) - z|$  instead of  $|x_i(k+1)|_{X_0}$ , following similar procedures with (7), (8), (9), and substituting (10) with (14), we can show that for any  $i \in \mathcal{V}$ ,

$$\begin{aligned} |x_i(k+1) - z| &\leq \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) |x_j(k) - z| + \tan \theta_k \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \alpha_{j,k} |x_j(k) - z| \\ &\quad - \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \alpha_{j,k} \left( |x_j(k) - z| - \sqrt{|x_j(k) - z|^2 - |x_j(k)|_{X_j}^2} \right). \end{aligned} \quad (15)$$

By dropping the non positive term on the righthand in (15), we have

$$\max_{1 \leq i \leq n} |x_i(k+1) - z| \leq (1 + \alpha_k^+ \tan \theta_k) \max_{1 \leq i \leq n} |x_i(k) - z|. \quad (16)$$

Therefore, for all  $k$ , we have

$$\begin{aligned} \max_{1 \leq i \leq n} |x_i(k+1) - z| &\leq \prod_{l=0}^k (1 + \alpha_l^+ \tan \theta_l) \max_{1 \leq i \leq n} |x_i(0) - z| \\ &\leq e^{\sum_{l=0}^k \alpha_l^+ \tan \theta_l} \max_{1 \leq i \leq n} |x_i(0) - z| \\ &\leq e^{\sum_{l=0}^{\infty} \alpha_l^+ \tan \theta_l} \max_{1 \leq i \leq n} |x_i(0) - z|, \end{aligned} \quad (17)$$

where the second inequality follows from  $1 + z \leq e^z$  for  $z \geq 0$ . Then the conclusion follows.  $\square$

The next lemma is a special case of various random versions, for example, see Lemma 11 in [6] (pp. 50). Here we give a simple proof.

**Lemma 4.3** *Let  $\{a_k\}_{k=0}^{\infty}$  and  $\{b_k\}_{k=0}^{\infty}$  be non-negative sequences with  $\sum_{k=0}^{\infty} b_k < \infty$ . Suppose*

$$a_{k+1} \leq a_k + b_k \text{ for all } k$$

*Then  $\lim_{k \rightarrow \infty} a_k$  is a finite number.*

Proof. Define a new sequence  $\{c_k\}_{k=1}^{\infty}$  with  $c_k = a_k - \sum_{l=0}^{k-1} b_l$ , which is bounded since  $\{a_k\}_{k=1}^{\infty}$  and  $\{\sum_{l=0}^{k-1} b_l\}_{k=1}^{\infty}$  are bounded. Moreover,  $c_{k+1} \leq c_k$  for  $k \geq 1$  implies that  $\lim_{k \rightarrow \infty} c_k$  is a finite number. The conclusion follows from  $\lim_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} c_k + \sum_{l=0}^{\infty} b_l$ .  $\square$

**Lemma 4.4** *Suppose A1 and A4 hold. If  $\sum_{k=0}^{\infty} \alpha_k^+ \theta_k < \infty$ , the following limit exists*

$$\vartheta := \lim_{k \rightarrow \infty} \max_{1 \leq i \leq n} |x_i(k)|_{X_0}.$$

Proof. Take  $z \in X_0$ . Based on (12) and (17), we have

$$\max_{1 \leq i \leq n} |x_i(k+1)|_{X_0} \leq \max_{1 \leq i \leq n} |x_i(k)|_{X_0} + \alpha_k^+ \tan \theta_k e^{\sum_{l=0}^{\infty} \alpha_l^+ \tan \theta_l} \max_{1 \leq i \leq n} |x_i(0) - z|.$$

The conclusion follows from the last inequality and Lemma 4.3.  $\square$

Denote

$$\eta_i^+ = \limsup_{k \rightarrow \infty} |x_i(k)|_{X_0}, \quad \eta_i^- = \liminf_{k \rightarrow \infty} |x_i(k)|_{X_0}, \quad i \in \mathcal{V}.$$

Obviously,  $0 \leq \eta_i^- \leq \eta_i^+ \leq \vartheta$  for all  $i$ .

**Lemma 4.5** Suppose **A1**–**A4** hold. If  $\sum_{k=0}^{\infty} \alpha_k^+ \theta_k < \infty$  and there exists some agent  $i_0 \in \mathcal{V}$  such that  $\eta_{i_0}^- < \vartheta$ , then  $\vartheta = 0$ .

Proof: Motivated by the idea of Lemma 4.3 in [31], we prove this lemma by contradiction.

Denote

$$\varrho_i = (\eta_i^- + \eta_i^+)/2, \quad i \in \mathcal{V}.$$

Since  $\eta_{i_0}^- < \vartheta$ , there exists an increasing sequence  $\{k_l\}_{l=0}^{\infty}$  such that  $|x_{i_0}(k_l)|_{X_0} \leq \varrho_{i_0} < \vartheta$  for  $l \geq 0$ . Moreover, for any  $\varepsilon > 0$ , there exist  $K_0 = K_0(\varepsilon)$  such that  $|x_i(k)|_{X_0} \leq \vartheta + \varepsilon$  and  $d_0 \sum_{k=K_0}^{\infty} \alpha_k^+ \theta_k \leq \varepsilon$  for  $k \geq K_0$  and all  $i$ , where

$$d_0 = (\tan \theta^* / \theta^*) \sup_{1 \leq i \leq n, k \geq 0} |x_i(k)|_{X_0}, \quad (18)$$

which is finite by Lemma 4.2. Without loss of generality, we assume  $k_0 \geq K_0$ .

Based on inequality (12), we have

$$|x_{i_0}(k_0 + 1)|_{X_0} \leq \sum_{j \in \mathcal{N}_{i_0}(k_0) \setminus i_0} a_{i_0 j}(k_0) |x_j(k_0)|_{X_0} + a_{i_0 i_0}(k_0) |x_{i_0}(k_0)|_{X_0} + d_0 \alpha_{k_0}^+ \theta_{k_0}.$$

Therefore,  $|x_{i_0}(k_0 + 1)|_{X_0} \leq (1 - \eta)(\vartheta + \varepsilon) + \eta \varrho_{i_0} + d_0 \alpha_{k_0}^+ \theta_{k_0}$  and then

$$\begin{aligned} |x_{i_0}(k_0 + 2)|_{X_0} &\leq (1 - \eta)(\vartheta + \varepsilon) + \eta[(1 - \eta)(\vartheta + \varepsilon) + d_0 \alpha_{k_0}^+ \theta_{k_0}] + \eta \varrho_{i_0} + d_0 \alpha_{k_0+1}^+ \theta_{k_0+1} \\ &\leq (1 - \eta^2)(\vartheta + \varepsilon) + \eta^2 \varrho_{i_0} + d_0 \sum_{k=k_0}^{k_0+1} \alpha_k^+ \theta_k. \end{aligned}$$

Similarly, we can show by induction that for  $r \geq 1$ ,

$$|x_{i_0}(k_0 + r)|_{X_0} \leq (1 - \eta^r)(\vartheta + \varepsilon) + \eta^r \varrho_{i_0} + d_0 \sum_{k=k_0}^{k_0+r-1} \alpha_k^+ \theta_k.$$

Since the communication graph is UJSC, there exist agent  $i_1 \neq i_0$  and time  $k_0^1 \in [k_0, k_0 + T)$  such that  $(i_1, i_0) \in \mathcal{E}(k_0^1)$ . As the above estimate for  $|x_{i_0}(k_0 + r)|_{X_0}$  with  $x_{i_0}(k_0)$ , by considering  $|x_{i_1}(k_0^1 + r)|_{X_0}$  with  $|x_{i_0}(k_0^1)|_{X_0}$ , we can show similarly that for  $r \geq 1$ ,

$$\begin{aligned} |x_{i_1}(k_0^1 + r)|_{X_0} &\leq (1 - \eta^r)(\vartheta + \varepsilon) + \eta^r |x_{i_0}(k_0^1)|_{X_0} + d_0 \sum_{k=k_0^1}^{k_0^1+r-1} \alpha_k^+ \theta_k \\ &\leq (1 - \eta^{k_0^1 - k_0 + r})(\vartheta + \varepsilon) + \eta^{k_0^1 - k_0 + r} \varrho_{i_0} + d_0 \sum_{k=k_0}^{k_0^1+r-1} \alpha_k^+ \theta_k. \end{aligned}$$

Repeating the previous procedure on intervals  $[k_0 + pT, k_0 + (p+1)T)$ ,  $1 \leq p \leq n-2$ , we can obtain nodes  $\{i_2, i_3, \dots, i_{n-1}\}$  such that  $\{i_j, 0 \leq j \leq n-1\} = \mathcal{V}$  and

$$\begin{aligned} \max_{1 \leq i \leq n} |x_i(k_0 + \hat{T})|_{X_0} &\leq (1 - \eta^{\hat{T}})(\vartheta + \varepsilon) + \eta^{\hat{T}} \varrho_{i_0} + d_0 \sum_{k=k_0}^{\infty} \alpha_k^+ \theta_k \\ &\leq (1 - \eta^{\hat{T}})(\vartheta + \varepsilon) + \eta^{\hat{T}} \varrho_{i_0} + \varepsilon, \end{aligned}$$

where  $\hat{T} = (n-1)T$ . Moreover, we can make similar analysis for  $k_1, k_2, \dots$  and obtain that for  $l \geq 0$ ,

$$\max_{1 \leq i \leq n} |x_i(k_l + \hat{T})|_{X_0} \leq (1 - \eta^{\hat{T}})(\vartheta + \varepsilon) + \eta^{\hat{T}} \varrho_{i_0} + \varepsilon,$$

which yields a contradiction since  $(1 - \eta^{\hat{T}})(\vartheta + \varepsilon) + \eta^{\hat{T}} \varrho_{i_0} + \varepsilon < \vartheta$  provided that  $\varepsilon$  is sufficiently small.  $\square$

We introduce transition matrices

$$\Phi(k, s) = A(k) \cdots A(s+1)A(s) \text{ for all } k \text{ and } s \text{ with } k \geq s.$$

Recall that  $\eta$  and  $T$  are defined in **A2** and **A3**, respectively and  $\hat{T} = (n-1)T$ . The next lemma generalizes Lemma 2 in [26] on the lower bound of the entries of the transition matrices.

**Lemma 4.6** *Suppose **A2** and **A3** hold. Then  $\Phi(k, s)_{ij} \geq \eta^{\hat{T}}$  for all  $i, j, s$  and  $k \geq s + \hat{T} - 1$ .*

Proof: By Lemma 2 in [26],  $\Phi(s + \hat{T} - 1, s)_{ij} \geq \eta^{\hat{T}}$  for all  $i, j$  and  $s \geq 0$ . Moreover, according to **A2** (i),  $\sum_{l=1}^n A(k)_{il} = 1$  and then  $\sum_{l=1}^n \Phi(k, s + \hat{T})_{il} = 1$  for all  $i, k$  and  $s$ . Thus, for all  $i, j$  and  $k \geq s + \hat{T} - 1$ ,

$$\begin{aligned} \Phi(k, s)_{ij} &= \left( \Phi(k, s + \hat{T}) \Phi(s + \hat{T} - 1, s) \right)_{ij} \\ &\geq \sum_{l=1}^n \Phi(k, s + \hat{T})_{il} \min_{1 \leq p, q \leq n} \Phi(s + \hat{T} - 1, s)_{pq} \\ &= \eta^{\hat{T}}. \end{aligned}$$

The conclusion follows.  $\square$

**Lemma 4.7**

$$\frac{1}{n} \sum_{i=1}^n \sqrt{v_0^2 - v_i^2} \leq \sqrt{v_0^2 - \left( \frac{\sum_{i=1}^n v_i}{n} \right)^2},$$

where  $v_0 \geq v_i \geq 0$  for all  $i$ .

Proof: The conclusion follows from that  $f(z) = \sqrt{c^2 - z^2}$  with domain  $[-c, c]$  is a concave function, where  $c > 0$ .  $\square$

Consider the following consensus model with noise  $w_i$ ,

$$z_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} b_{ij}(k) z_j(k) + w_i(k), \quad i = 1, \dots, n, \quad (19)$$

where the weights  $b_{ij}(k), i, j \in \mathcal{V}, k \geq 0$  satisfy **A2**. The consensus is said to be achieved for system (19) if for any initial conditions,  $\lim_{k \rightarrow \infty} |z_i(k) - z_j(k)| = 0$  for all  $1 \leq i, j \leq n$ . The next lemma can be obtained from Theorem 1 in [34].

**Lemma 4.8** *If the communication graph of system (19) is UJSC with  $\lim_{k \rightarrow \infty} w_i(k) = 0$  for all  $i$ , then the consensus is achieved for system (19).*

## 4.2 Proofs

In this subsection, we present the proofs of Theorems 4.1 and 4.2.

### 4.2.1 Proof of Theorem 4.1

Rewrite (5) as

$$\begin{aligned} x_i(k+1) = & \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)x_j(k) + \left( \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}(P_{X_j}(x_j(k)) - x_j(k)) \right) \\ & + \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)\alpha_{j,k}\left(P_j^{sa}(k) - P_{X_j}(x_j(k))\right). \end{aligned} \quad (20)$$

Based on (9), the second and the third term in (20) are not greater than

$$\max_{1 \leq i \leq n} \alpha_{i,k}|x_i(k)|_{X_i} + \alpha_k^+ \tan \theta_k \max_{1 \leq i \leq n} |x_i(k)|_{X_i}. \quad (21)$$

Note that  $\vartheta = 0$  leads to  $\lim_{k \rightarrow \infty} \max_{1 \leq i \leq n} |x_i(k)|_{X_i} \leq \lim_{k \rightarrow \infty} \max_{1 \leq i \leq n} |x_i(k)|_{X_0} = 0$  and then the term in (21) tends to zero as  $k \rightarrow \infty$ . Therefore, by applying Lemma 4.8 for (20), we have that if  $\vartheta = 0$ , then the consensus is achieved.

Moreover, we claim that if  $\vartheta = 0$  and the consensus is achieved, then all agents will converge to a point in  $X_0$ . Since  $\{x_i(k)\}_{k=0}^\infty, i = 1, \dots, n$  are bounded by Lemma 4.2 and the consensus is achieved, there is  $x^* \in X_0$  and a subsequence  $\{k_l\}_{l=1}^\infty$  such that  $\lim_{l \rightarrow \infty} x_i(k_l) = x^*$  for all  $i$ . By similar analysis with (17), we have

$$\max_{1 \leq i \leq n} |x_i(k) - x^*| \leq e^{\sum_{p=k_l}^\infty \alpha_p^+ \tan \theta_p} \max_{1 \leq i \leq n} |x_i(k_l) - x^*| \leq e^{\sum_{p=0}^\infty \alpha_p^+ \tan \theta_p} \max_{1 \leq i \leq n} |x_i(k_l) - x^*|$$

for  $k \geq k_l$ , which implies  $\lim_{k \rightarrow \infty} x_i(k) = x^*$  for all  $i$ .

If there exists some agent  $i_0$  such that  $\eta_{i_0}^- < \vartheta$ , then by Lemma 4.5,  $\vartheta = 0$ . Therefore, we only need to prove

$$\vartheta = 0 \text{ when } \eta_i^+ = \eta_i^- = \vartheta \text{ for all } i,$$

which shall be proven by contradiction. If  $\vartheta > 0$ , then for any  $\varepsilon > 0$ , there exist  $K_1 = K_1(\varepsilon)$  such that  $|x_i(k)|_{X_0} \leq \vartheta + \varepsilon$  and  $d_0 \alpha_k^+ \theta_k \leq \varepsilon$  for  $k \geq K_1$  and all  $i$ . We complete the proof by the following two steps.

- (i) Suppose  $\eta_i^+ = \eta_i^- = \vartheta$  for all  $i$ . The consensus is achieved:  $\lim_{k \rightarrow \infty} |x_i(k) - x_j(k)| = 0$  for all  $i, j$ .

Denote

$$\varsigma_i = \limsup_{k \rightarrow \infty} \alpha_{i,k}|x_i(k)|_{X_i}, \quad i \in \mathcal{V}.$$

We prove  $\varsigma_i = 0$  for all  $i$  by contradiction. If there exists some agent  $i_0$  such that  $\varsigma_{i_0} > 0$ , then there is an increasing time subsequence  $\{k_l\}_{l=1}^\infty$  with  $k_1 \geq K_1$  such that

$\alpha_{i_0, k_l} |x_{i_0}(k_l)|_{X_{i_0}} \geq c \varsigma_{i_0}$  for all  $l$  and some  $0 < c < 1$ . Therefore, by (11)

$$\begin{aligned}
|x_{i_0}(k_l + 1)|_{X_0} &\leq a_{i_0 i_0}(k) \left( (1 - \alpha_{i_0, k_l}) |x_{i_0}(k_l)|_{X_0} + \alpha_{i_0, k_l} \sqrt{|x_{i_0}(k_l)|_{X_0}^2 - |x_{i_0}(k_l)|_{X_{i_0}}^2} \right) \\
&\quad + \sum_{j \in \mathcal{N}_{i_0}(k_l) \setminus i_0} a_{i_0 j}(k_l) |x_j(k_l)|_{X_0} + d_0 \alpha_{k_l}^+ \theta_{k_l} \\
&\leq \eta \left( (1 - \alpha_{i_0, k_l}) (\vartheta + \varepsilon) + \sqrt{\alpha_{i_0, k_l}^2 (\vartheta + \varepsilon)^2 - c^2 \varsigma_{i_0}^2} \right) + (1 - \eta) (\vartheta + \varepsilon) + \varepsilon \\
&= (1 - \eta \alpha_{i_0, k_l}) (\vartheta + \varepsilon) + \eta \sqrt{\alpha_{i_0, k_l}^2 (\vartheta + \varepsilon)^2 - c^2 \varsigma_{i_0}^2} + \varepsilon, \tag{22}
\end{aligned}$$

where  $d_0$  is the one in (18), which yields a contradiction since the right hand side of (22) is less than  $\vartheta$  for sufficiently small  $\varepsilon$  and sufficiently large  $l$ .

Thus,  $\lim_{k \rightarrow \infty} \alpha_{i, k} |x_i(k)|_{X_i} = 0$  for all  $i$ . Moreover, since  $\sum_{k=0}^{\infty} \alpha_k^+ \theta_k < \infty$ ,  $\lim_{k \rightarrow \infty} \alpha_k^+ \tan \theta_k \leq (\tan \theta^* / \theta^*) \lim_{k \rightarrow \infty} \alpha_k^+ \theta_k = 0$ . The two preceding conclusions and the boundedness of  $\{x_i(k)\}_{k=0}^{\infty}$  imply that the term in (21) tends to zero and then the consensus is achieved by applying Lemma 4.8 for (20) again.

- (ii) Suppose  $\eta_i^+ = \eta_i^- = \vartheta$  for all  $i$ . All agents converge to the nonempty intersection set:  $\lim_{k \rightarrow \infty} |x_i(k)|_{X_0} = 0$  for all  $i$ .

Denote

$$\delta = \liminf_{k \rightarrow \infty} \sum_{i=1}^n |x_i(k)|_{X_i}.$$

We prove that  $\delta = 0$  by contradiction. Otherwise, suppose  $\delta > 0$ .

By (6), we obtain for  $k \geq s$ ,

$$\begin{aligned}
|x(k+1)|_{X_0} &\leq \Phi(k, s) |x(s)|_{X_0} - \sum_{l=s}^k \Phi(k, l) D_l y(l) + d_0 \sum_{l=s}^k \alpha_l^+ \theta_l \\
&= \Phi(k, s) |x(s)|_{X_0} - \sum_{l=s}^{k-\hat{T}+1} \Phi(k, l) D_l y(l) \\
&\quad - \sum_{l=k-\hat{T}+2}^k \Phi(k, l) D_l y(l) + d_0 \sum_{l=s}^k \alpha_l^+ \theta_l, \tag{23}
\end{aligned}$$

where  $\hat{T} = (n-1)T$ . By dropping the third term (nonpositive) on the right-hand side in (23), we obtain

$$|x(k+1)|_{X_0} \leq \Phi(k, s) |x(s)|_{X_0} - \sum_{l=s}^{k-\hat{T}+1} \Phi(k, l) D_l y(l) + d_0 \sum_{l=s}^k \alpha_l^+ \theta_l. \tag{24}$$

For  $\bar{\varepsilon} = \delta^2 / (4n^2 \vartheta + 2\delta)$ , there exists sufficiently large  $K_2$  such that  $\sum_{i=1}^n |x_i(k)|_{X_i} > \delta - \bar{\varepsilon}$

and  $\vartheta - \bar{\varepsilon} \leq |x_i(k)|_{X_0} \leq \vartheta + \bar{\varepsilon}$  for  $k \geq K_2$ . For  $k \geq K_2$ , we have

$$\begin{aligned} \sum_{i=1}^n \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} &\leq \sum_{i=1}^n \sqrt{(\vartheta + \bar{\varepsilon})^2 - |x_i(k)|_{X_i}^2} \\ &\leq n \sqrt{(\vartheta + \bar{\varepsilon})^2 - \left( \left( \sum_{i=1}^n |x_i(k)|_{X_i} \right) / n \right)^2} \\ &\leq n \sqrt{(\vartheta + \bar{\varepsilon})^2 - ((\delta - \bar{\varepsilon})/n)^2}, \end{aligned}$$

where the second inequality follows from Lemma 4.7 and then

$$\begin{aligned} \sum_{i=1}^n \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right) &\geq n \left( \vartheta - \bar{\varepsilon} - \sqrt{(\vartheta + \bar{\varepsilon})^2 - ((\delta - \bar{\varepsilon})/n)^2} \right) \\ &:= \zeta > 0. \end{aligned}$$

Namely,  $\sum_{i=1}^n y_i(l) \geq \zeta$  for  $l \geq K_2$ . Combining the preceding inequality with Lemma 4.6, we have that every component of  $\Phi(k, l)D_l y(l)$  is not less than  $\eta^{\hat{T}} \zeta \alpha_l^-$  for all  $K_2 \leq l \leq k - \hat{T} + 1$  and all  $k \geq K_2 + \hat{T} - 1$ . Then by (24) with taking  $s = K_2$ , we obtain

$$|x(k+1)|_{X_0} \leq \Phi(k, K_2)|x(K_2)|_{X_0} - \eta^{\hat{T}} \zeta \sum_{l=K_2}^{k-\hat{T}+1} \alpha_l^- \mathbf{1} + d_0 \sum_{l=K_2}^k \alpha_l^+ \theta_l, \quad (25)$$

where  $\mathbf{1}$  is the vector of all ones. Observing that  $\sum_{l=K_2}^{\infty} \alpha_l^- = \infty$ ,  $\sum_{l=K_2}^{\infty} \alpha_l^+ \theta_l < \infty$ , and noticing  $\lim_{k \rightarrow \infty} |x(k)|_{X_0} = \vartheta \mathbf{1}$ , a contradiction arises by taking the limit as  $k \rightarrow \infty$  in (25).

Therefore,  $\delta = 0$ , that is, there is a subsequence  $\{k_l\}_{l=0}^{\infty}$  such that  $\lim_{l \rightarrow \infty} \sum_{i=1}^n |x_i(k_l)|_{X_i} = 0$ . Since the consensus is achieved by what we have proven in the first step (i), we have

$$\lim_{l \rightarrow \infty} \sum_{i=1}^n |x_i(k_l)|_{X_j} = 0 \text{ for all } j \in \mathcal{V},$$

which implies  $\vartheta = \lim_{l \rightarrow \infty} \max_{1 \leq i \leq n} |x_i(k_l)|_{X_0} = 0$ .

This completes the proof.  $\square$

**Remark 4.3** If the adjacency matrices are double-stochastic,  $\sum_{k=0}^{\infty} \alpha_k^- = \infty$  and  $\sum_{k=0}^{\infty} \alpha_k^+ \theta_k < \infty$ , then the convergence analysis for the optimal consensus can also be performed by the following procedures.

By summing the two sides in (12) over  $i = 1, \dots, n$ , we obtain

$$\sum_{i=1}^n |x_i(k+1)|_{X_0} \leq (1 + \alpha_k^+ \tan \theta_k) \sum_{i=1}^n |x_i(k)|_{X_0} - \sum_{i=1}^n \alpha_{i,k} \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right). \quad (26)$$

Summing the two sides in (26) over  $k \geq 0$  and rearranging the terms, we have

$$\begin{aligned}
\sum_{k=0}^{\infty} \alpha_k^- \sum_{i=1}^n \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right) \\
\leq \sum_{k=0}^{\infty} \sum_{i=1}^n \alpha_{i,k} \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right) \\
\leq \sum_{i=1}^n |x_i(0)|_{X_0} + \sum_{k=0}^{\infty} \alpha_k^+ \tan \theta_k \sum_{i=1}^n |x_i(k)|_{X_0} < \infty.
\end{aligned} \tag{27}$$

The assumption  $\sum_{k=0}^{\infty} \alpha_k^- = \infty$  and (27) imply that

$$\liminf_{k \rightarrow \infty} \sum_{i=1}^n \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right) = 0,$$

and then

$$\liminf_{k \rightarrow \infty} \sum_{i=1}^n |x_i(k)|_{X_i} = 0. \tag{28}$$

From (27) we also have

$$\lim_{k \rightarrow \infty} \sum_{i=1}^n \alpha_{i,k} \left( |x_i(k)|_{X_0} - \sqrt{|x_i(k)|_{X_0}^2 - |x_i(k)|_{X_i}^2} \right) = 0 \tag{29}$$

and then for all  $i \in \mathcal{V}$ ,

$$\limsup_{k \rightarrow \infty} \alpha_{i,k} |x_i(k)|_{X_i} = 0 \tag{30}$$

since if there is  $i_0 \in \mathcal{V}$  and  $\{k_l\}_{l=1}^{\infty}$  such that  $\alpha_{i_0, k_l} |x_{i_0}(k_l)|_{X_{i_0}} \geq \varepsilon$  for all  $l$  and some  $\varepsilon > 0$ , then for all  $l$ ,

$$\alpha_{i_0, k_l} \left( |x_{i_0}(k_l)|_{X_0} - \sqrt{|x_{i_0}(k_l)|_{X_0}^2 - |x_{i_0}(k_l)|_{X_{i_0}}^2} \right) \geq \alpha_{i_0, k_l} |x_{i_0}(k_l)|_{X_0} - \sqrt{\alpha_{i_0, k_l}^2 |x_{i_0}(k_l)|_{X_0}^2 - \varepsilon^2} > 0,$$

which contradicts with (29). As what have been proved in Theorem 4.1, (28) and (30) imply that the optimal consensus is achieved.

#### 4.2.2 Proof of Theorem 4.2

It is easy to find that if  $\theta_k \equiv 0$ , the intersection set in (1) is the line segment from  $x_i(k)$  to  $P_{X_i}(x_i(k))$  and then  $P_i^{sa}(k) = P_{X_i}(x_i(k))$ . Then the evolution of the approximate projected consensus algorithm becomes

$$x_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \left( (1 - \alpha_{j,k}) x_j(k) + \alpha_{j,k} P_{X_j}(x_j(k)) \right), \quad i = 1, \dots, n. \tag{31}$$

We complete the proof by the following two parts.

- (i) We first prove that if  $\sum_{k=0}^{\infty} \alpha_k^+ < \infty$ , then there exist initial conditions from which all agents will not converge to  $X_0$ . Let  $z^* \in R^m$ , which shall be selected later, and  $x_i(0) = z^*, i = 1, \dots, n$ .

By (31),  $x_i(1)$  can be rewritten as

$$\begin{aligned} x_i(1) &= \sum_{j \in \mathcal{N}_i(0)} a_{ij}(0) \left( (1 - \alpha_{j,0})x_j(0) + \alpha_{j,0}P_{X_j}(x_j(0)) \right) \\ &= \sum_{j \in \mathcal{N}_i(0)} a_{ij}(0)(1 - \alpha_{j,0})z^* + \sum_{j \in \mathcal{N}_i(0)} a_{ij}(0)\alpha_{j,0}P_{X_0}(z^*) + \Delta_{i0}, \\ &= (1 - \beta_{i,0})z^* + \beta_{i,0}P_{X_0}(z^*) + \Delta_{i0}, \end{aligned}$$

where  $1 - \beta_{i,0} = \sum_{j \in \mathcal{N}_i(0)} a_{ij}(0)(1 - \alpha_{j,0})$  and  $\Delta_{i0} = \sum_{j \in \mathcal{N}_i(0)} a_{ij}(0)\alpha_{j,0}(P_{X_j}(z^*) - P_{X_0}(z^*))$  with  $|\Delta_{i0}| \leq \alpha_0^+ d^*$  for all  $i$ .

We also have

$$\begin{aligned} x_i(2) &= \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1) \left( (1 - \alpha_{j,1})x_j(1) + \alpha_{j,1}P_{X_j}(x_j(1)) \right) \\ &= \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)(1 - \alpha_{j,1}) \left( (1 - \beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*) \right) + \Delta_{i1} \\ &\quad + \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1}P_{X_0} \left( (1 - \beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*) \right) \\ &= (1 - \beta_{i,1})z^* + \beta_{i,1}P_{X_0}(z^*) + \Delta_{i1}, \end{aligned}$$

where  $1 - \beta_{i,1} = \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)(1 - \alpha_{j,1})(1 - \beta_{j,0})$ , the third equality follows from Lemma 2.2 (iii) and  $\Delta_{i1} = \Delta_{i1}^1 + \Delta_{i1}^2 + \Delta_{i1}^3$  with

$$\begin{aligned} \Delta_{i1}^1 &= \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)(1 - \alpha_{j,1})\Delta_{j0}; \\ \Delta_{i1}^2 &= \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1} \left( P_{X_j}(x_j(1)) - P_{X_0}(x_j(1)) \right); \\ \Delta_{i1}^3 &= \sum_{j \in \mathcal{N}_i(1)} a_{ij}(1)\alpha_{j,1} \left( P_{X_0}(x_j(1)) - P_{X_0}((1 - \beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*)) \right). \end{aligned} \quad (32)$$

Now we give an estimate of the upper bound of  $\Delta_{i1}$ . By Lemma 2.2 (i),

$$\begin{aligned} |P_{X_0}(x_j(1)) - P_{X_0}((1 - \beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*))| &\leq |x_j(1) - ((1 - \beta_{j,0})z^* + \beta_{j,0}P_{X_0}(z^*))| \\ &= |\Delta_{j0}|, \end{aligned}$$

which implies that  $|\Delta_{i1}^1| + |\Delta_{i1}^3| \leq \max_{1 \leq i \leq n} |\Delta_{i0}| \leq \alpha_0^+ d^*$  and then  $|\Delta_{i1}| \leq |\Delta_{i1}^1| + |\Delta_{i1}^3| + |\Delta_{i1}^2| \leq (\alpha_0^+ + \alpha_1^+)d^*$  for all  $i$ .

Similarly, we can show by induction that for all  $i$  and  $k$ ,  $x_i(k+1)$  can be expressed as

$$x_i(k+1) = (1 - \beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*) + \Delta_{ik}, \quad (33)$$



where  $|\Delta_{ik}| \leq \sum_{l=0}^k \alpha_l^+ d^*$  and  $\{\beta_{i,k}, i \in \mathcal{V}\}_{k=0}^\infty$  satisfies

$$1 - \beta_{i,k} = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)(1 - \alpha_{j,k})(1 - \beta_{j,k-1}). \quad (34)$$

Based on (34), we can show by induction that

$$1 - \beta_{i,k} \geq \prod_{l=0}^k (1 - \alpha_l^+) \text{ for all } i \text{ and } k. \quad (35)$$

It follows from (33) and Lemma 2.2 (ii) that

$$|x_i(k+1)|_{X_0} \geq |(1 - \beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*)|_{X_0} - |\Delta_{ik}|. \quad (36)$$

Moreover, for a convex set  $K$  and any  $0 \leq \lambda \leq 1$ ,

$$\begin{aligned} |(1 - \lambda)x + \lambda P_K(x)|_K &= |(1 - \lambda)x + \lambda P_K(x) - P_K((1 - \lambda)x + \lambda P_K(x))| \\ &= |(1 - \lambda)x + \lambda P_K(x) - P_K(x)| \\ &= (1 - \lambda)|x|_K, \end{aligned} \quad (37)$$

where the second equality follows from Lemma 2.2 (iii). By taking  $\lambda = \beta_{i,k}$  in (37), we have

$$|(1 - \beta_{i,k})z^* + \beta_{i,k}P_{X_0}(z^*)|_{X_0} = (1 - \beta_{i,k})|z^*|_{X_0}.$$

Combining the last equality, (36) and (35), we obtain

$$|x_i(k+1)|_{X_0} \geq \prod_{l=0}^k (1 - \alpha_l^+) |z^*|_{X_0} - \sum_{l=0}^k \alpha_l^+ d^*. \quad (38)$$

Taking the inferior limit on the two sides in (38), we have for all  $i$ ,

$$\liminf_{k \rightarrow \infty} |x_i(k)|_{X_0} \geq \prod_{l=0}^{\infty} (1 - \alpha_l^+) |z^*|_{X_0} - \sum_{l=0}^{\infty} \alpha_l^+ d^*,$$

which is positive provided that

$$|z^*|_{X_0} > \frac{\sum_{l=0}^{\infty} \alpha_l^+ d^*}{\prod_{l=0}^{\infty} (1 - \alpha_l^+)}, \quad (39)$$

where  $\prod_{l=0}^{\infty} (1 - \alpha_l^+) > 0$  since  $0 \leq \alpha_l^+ < 1$  for all  $l$  and  $\sum_{l=0}^{\infty} \alpha_l^+ < \infty$ . Thus, all agents can not achieve an optimal consensus for all initial conditions satisfying (39).

- (ii) In fact, if  $\sum_{k=0}^{\infty} \alpha_k^+ < \infty$ , there is  $y^* = y^*(z^*) \notin X_0$  such that  $\lim_{k \rightarrow \infty} x_i(k) = y^*$  for all  $i$  provided  $z^*$  satisfies (39).

Denote  $\bar{d} = \sup_{1 \leq i \leq n, k \geq 0} |x_i(k)|_{X_i}$ . By (31), we have for all  $i$  and  $k$ ,

$$x_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k)x_j(k) + \Gamma_{i,k}, \quad (40)$$

where  $\Gamma_{i,k} = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) \alpha_{jk}(P_{X_j}(x_j(k)) - x_j(k))$  with  $|\Gamma_{i,k}| \leq \bar{d} \alpha_k^+$ .

Take  $i_0 \in \mathcal{V}$ . It follows from (40) that for all  $j$  and  $k$ ,

$$|x_j(k+1) - x_{i_0}(k)| \leq \max_{1 \leq p, q \leq n} |x_p(k) - x_q(k)| + \max_{1 \leq r \leq n} |\Gamma_{r,k}|. \quad (41)$$

Then

$$\begin{aligned} & |x_{i_0}(k+2) - x_{i_0}(k)| \\ &= \left| \sum_{j \in \mathcal{N}_{i_0}(k+1)} a_{i_0j}(k+1) x_j(k+1) + \sum_{j \in \mathcal{N}_{i_0}(k+1)} a_{i_0j}(k+1) \Gamma_{j,k+1} - x_{i_0}(k) \right| \\ &\leq \max_{1 \leq j \leq n} |x_j(k+1) - x_{i_0}(k)| + \max_{1 \leq r \leq n} |\Gamma_{r,k+1}| \\ &\leq \max_{1 \leq p, q \leq n} |x_p(k) - x_q(k)| + \max_{1 \leq r \leq n} |\Gamma_{r,k}| + \max_{1 \leq r \leq n} |\Gamma_{r,k+1}| \\ &\leq \max_{1 \leq p, q \leq n} |x_p(k) - x_q(k)| + \bar{d}(\alpha_k^+ + \alpha_{k+1}^+). \end{aligned}$$

We can similarly show by induction that for all  $k$  and  $l$ ,

$$\begin{aligned} |x_{i_0}(k+l) - x_{i_0}(k)| &\leq \max_{1 \leq p, q \leq n} |x_p(k) - x_q(k)| + \bar{d} \sum_{s=k}^{k+l-1} \alpha_s^+ \\ &\leq \max_{1 \leq p, q \leq n} |x_p(k) - x_q(k)| + \bar{d} \sum_{s=k}^{\infty} \alpha_s^+. \end{aligned} \quad (42)$$

Since  $\lim_{k \rightarrow \infty} \alpha_k^+ = 0$  and  $\{x_i(k)\}_{k=0}^{\infty}$  is bounded for all  $i$ , the consensus is achieved by Lemma 4.8. Combining the consensus and the boundedness of  $\{x_i(k)\}_{k=0}^{\infty}$ , there exist a subsequence  $\{k_l\}_{l=0}^{\infty}$  and  $y^*$  such that  $\lim_{l \rightarrow \infty} x_i(k_l) = y^*$  for all  $i$ . Therefore, since  $\sum_{k=0}^{\infty} \alpha_k^+ < \infty$ , for any  $\varepsilon > 0$ , there exists  $l_0 = l_0(\varepsilon)$  such that  $|x_i(k_l) - x_j(k_l)| \leq \varepsilon/2$  and  $\bar{d} \sum_{s=k_l}^{\infty} \alpha_s^+ \leq \varepsilon/2$  for all  $i, j$  and  $l \geq l_0$ . Thus, for  $l \geq l_0$  and  $k \geq k_l$ ,

$$|x_{i_0}(k) - x_{i_0}(k_l)| \leq \max_{1 \leq p, q \leq n} |x_p(k_l) - x_q(k_l)| + \bar{d} \sum_{s=k_l}^{\infty} \alpha_s^+ \leq \varepsilon,$$

which implies that  $\lim_{l \rightarrow \infty} x_{i_0}(k) = y^*$  since  $\varepsilon$  can be arbitrarily small. By what we have proven in the first part (i),  $y^* \notin X_0$ . The desired conclusion follows from the fact that  $i_0$  is taken arbitrarily.

The proof is completed.  $\square$

## 5 Critical Angle Error

The boundedness of system states plays a key role for various optimization methods [28, 26]. In this section, we consider the effect of the angle error  $\theta_k$  on the boundedness of the states  $\{x_i(k), i \in \mathcal{V}\}_{k=0}^{\infty}$  generated by the approximate projected consensus algorithm.

Suppose  $\alpha_{i,k} \equiv 1$  and  $\theta_k \equiv \theta$  with  $0 < \theta < \pi/2$ . First the following conclusion shows that when  $\theta < \pi/4$ , the trajectories of the algorithm are uniformly bounded with respect to all initial conditions.

**Proposition 5.1** Suppose **A1**, **A2**, **A5** hold and  $0 < \theta < \pi/4$ . Then we have

$$\sup_{x(0)} \limsup_{k \rightarrow \infty} |x_i(k)|_{X_0} < \infty$$

for  $i = 1, \dots, n$ .

Proof: Recall  $d^* = \sup_{\omega_1, \omega_2 \in \bigcup_{i=1}^n X_i} |\omega_1 - \omega_2|$ . We claim that for all  $i$  and all initial conditions  $x(0)$ ,

$$\limsup_{k \rightarrow \infty} |x_i(k)|_{X_0} \leq \frac{2d^*}{1 - \tan \theta},$$

which implies the conclusion.

Based on (8), (9) and the definition of  $d^*$ , we always have

$$\begin{aligned} |P_i^{sa}(k)|_{X_0} &\leq \tan \theta |x_i(k)|_{X_i} + |P_{X_i}(x_i(k))|_{X_0} \\ &\leq \tan \theta |x_i(k)|_{X_0} + d^*, \end{aligned} \quad (43)$$

Furthermore, in the case of  $|x_i(k)|_{X_0} \geq 2d^*/(1 - \tan \theta)$ , by (43) we further obtain

$$\begin{aligned} |P_i^{sa}(k)|_{X_0} &\leq \tan \theta |x_i(k)|_{X_0} + d^* \\ &\leq \left(1 - \frac{1 - \tan \theta}{2}\right) |x_i(k)|_{X_0}. \end{aligned} \quad (44)$$

We next consider  $h(k) := \max_{1 \leq i \leq n} |x_i(k)|_{X_0}$  by the following two cases:

(i)  $h(k) \leq 2d^*/(1 - \tan \theta)$ . By Lemma 2.1 and (43), we have for all  $i$ ,

$$\begin{aligned} |x_i(k+1)|_{X_0} &\leq \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) |P_j^{sa}(k)|_{X_0} \\ &= \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) (\tan \theta |x_j(k)|_{X_0} + d^*) \\ &\leq \frac{2d^* \tan \theta}{1 - \tan \theta} + d^* \\ &\leq \frac{2d^*}{1 - \tan \theta}. \end{aligned} \quad (45)$$

Therefore,  $h(k+1) \leq 2d^*/(1 - \tan \theta)$ .

(ii)  $h(k) > 2d^*/(1 - \tan \theta)$ . Define  $\mathcal{V}_1(k) = \{i \mid |x_j(k)|_{X_0} \leq 2d^*/(1 - \tan \theta) \text{ for all } j \in \mathcal{N}_i(k)\}$  and  $\mathcal{V}_2(k) = \mathcal{V} \setminus \mathcal{V}_1(k)$ , where  $\mathcal{V}_2(k)$  is nonempty.

1) If  $i \in \mathcal{V}_1(k)$ , by similar analysis with (45), we obtain  $|x_i(k+1)|_{X_0} \leq 2d^*/(1 - \tan \theta)$ .

2) If  $i \in \mathcal{V}_2(k)$ , we can define  $\mathcal{N}_i^1(k) = \{j \mid j \in \mathcal{N}_i(k), |x_j(k)|_{X_0} \leq 2d^*/(1 - \tan \theta)\}$  and  $\mathcal{N}_i^2(k) = \mathcal{N}_i(k) \setminus \mathcal{N}_i^1(k)$ , where  $\mathcal{N}_i^2(k)$  is nonempty.

Again by Lemma 2.1, we have

$$|x_i(k+1)|_{X_0} \leq \sum_{j \in \mathcal{N}_i^1(k)} a_{ij}(k) |P_j^{sa}(k)|_{X_0} + \sum_{j \in \mathcal{N}_i^2(k)} a_{ij}(k) |P_j^{sa}(k)|_{X_0}. \quad (46)$$

For  $j \in \mathcal{N}_i^1(k)$ , by (43) we have  $|P_j^{sa}(k)|_{X_0} \leq \tan \theta |x_j(k)|_{X_0} + d^* \leq 2d^* \tan \theta / (1 - \tan \theta) + d^* \leq 2d^* / (1 - \tan \theta)$ . As a result, by (44) and (46) we have

$$|x_i(k+1)|_{X_0} \leq \frac{2d^*}{1 - \tan \theta} \sum_{j \in \mathcal{N}_i^1(k)} a_{ij}(k) + \left(1 - \frac{1 - \tan \theta}{2}\right) \sum_{j \in \mathcal{N}_i^2(k)} a_{ij}(k) |x_j(k)|_{X_0}. \quad (47)$$

Since  $|x_j(k)|_{X_0} \geq 2d^* / (1 - \tan \theta)$  for all  $j \in \mathcal{N}_i^2(k)$ ,  $\max_{j \in \mathcal{N}_i(k)} |x_j(k)|_{X_0} \geq 2d^* / (1 - \tan \theta)$ . Thus, it follows from (47) that

$$\begin{aligned} |x_i(k+1)|_{X_0} &\leq \left( \sum_{j \in \mathcal{N}_i^1(k)} a_{ij}(k) + \left(1 - \frac{1 - \tan \theta}{2}\right) \sum_{j \in \mathcal{N}_i^2(k)} a_{ij}(k) \right) \max_{j \in \mathcal{N}_i(k)} |x_j(k)|_{X_0} \\ &\leq \left( (1 - \eta) + \eta \left(1 - \frac{1 - \tan \theta}{2}\right) \right) h(k+1) \\ &= \left(1 - \frac{\eta(1 - \tan \theta)}{2}\right) h(k+1), \end{aligned}$$

where  $0 < 1 - \eta(1 - \tan \theta)/2 < 1$  and  $\eta$  is the lower bound of weights in **A2**.

Therefore, based on the two cases (1) and (2), we show that if  $h(k) > 2d^* / (1 - \tan \theta)$ , then

$$h(k+1) \leq \max \{2d^* / (1 - \tan \theta), (1 - \eta(1 - \tan \theta)/2) h(k)\}.$$

Combining the two cases (i) and (ii), we have

$$h(k+1) \leq \begin{cases} 2d^* / (1 - \tan \theta), & \text{if } h(k) \leq 2d^* / (1 - \tan \theta); \\ \max \{2d^* / (1 - \tan \theta), (1 - \eta(1 - \tan \theta)/2) h(k)\}, & \text{otherwise} \end{cases}$$

Thus, the conclusion follows.  $\square$

We next investigate the non-conservativeness of  $\pi/4$  in Proposition 5.1. We just focus on a special case with only one node in the network. Its set is denoted as  $X_*$ . We denote the states of the node as  $\{x_*(k)\}_{k=0}^\infty$  driven by the approximate projected consensus algorithm:

$$x_*(k+1) \in \mathcal{P}_{X_*}^{sa}(x_*(k), \theta),$$

where  $\mathcal{P}_{X_*}^{sa}(x_*(k), \theta) = \mathbf{C}_{X_*}(x_*(k), \theta) \cap \mathbf{H}_{X_*}(x_*(k))$ .

The following conclusion holds.

**Proposition 5.2** Suppose  $\theta = \pi/4$ .

(i) For any compact convex set  $X_*$  and any initial condition  $x_*(0) \in R^m$ , we have

$$\limsup_{k \rightarrow \infty} |x_*(k)|_{X_*} \leq |x_*(0)|_{X_*};$$

(ii) There exists an approximate projection sequence  $\{P_*^{sa}(k)\}_{k=0}^\infty$ ,  $P_*^{sa}(k) \in \mathcal{P}_{X_*}^{sa}(x_*(k), \pi/4)$  such that

(ii.1)  $\limsup_{k \rightarrow \infty} |x_*(k)|_{X_*} = 0$  when  $X_*$  is a ball with radius  $r > 0$ ;

(ii.2)  $\limsup_{k \rightarrow \infty} |x_*(k)|_{X_*} = |x_*(0)|_{X_*}$  when  $X_*$  is a single point.

Proof: (i) The conclusion follows from that

$$|x_*(k+1)|_{X_*} \leq |x_*(k+1) - P_{X_*}(x_*(k))| \leq \tan \theta |x_*(k)|_{X_*} = |x_*(k)|_{X_*}.$$

(ii) We just select  $\{P_*^{sa}(k)\}_{k=0}^\infty$  for which  $\text{Ang}(P_*^{sa}(k) - x_*(k), P_{X_*}(x_*(k)) - x_*(k)) = \pi/4$  for all  $k$  satisfying the definition (1) such that the solution of the algorithm with  $\alpha_{i,k} \equiv 1$  and  $\theta_k \equiv \pi/4$  satisfies the following two cases.

(ii.1) It is easy to find that for all  $k$ ,

$$|x_*(k+1)|_{X_*} = \sqrt{|x_*(k)|_{X_*}^2 + r^2} - r, \quad (48)$$

which implies that the sequence  $\{|x_*(k)|_{X_*}\}_{k=0}^\infty$  is non-increasing and then converges to some nonnegative number  $\mu$ . Letting  $k \rightarrow \infty$  in (48), we have  $\mu = \sqrt{\mu^2 + r^2} - r$  and then  $\mu = 0$ .

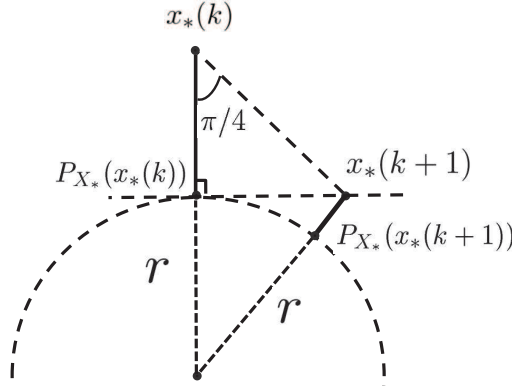


Figure 3: Approximate projection with respect to a ball.

(ii.2) The conclusion is straightforward.  $\square$

We present another result for the case when  $\theta > \pi/4$  to reveal that, in this case, the node states will be unbounded as long as the distance between the initial condition and  $X_*$  is larger than a certain threshold.

**Proposition 5.3** *Suppose  $\theta > \pi/4$ . Then for any compact convex set  $X_*$ , there exists an approximate projection sequence  $\{P_*^{sa}(k)\}_{k=0}^\infty$  such that*

$$\limsup_{k \rightarrow \infty} |x_*(k)|_{X_*} = \infty$$

for all initial conditions satisfying  $|x_*(0)|_{X_*} > \sup_{\omega_1, \omega_2 \in X_*} |\omega_1 - \omega_2| / (\tan \theta - 1)$ .

Proof: Select  $\{P_*^{sa}(k)\}_{k=0}^\infty$  for which  $\text{Ang}(P_*^{sa}(k) - x_*(k), P_{X_*}(x_*(k)) - x_*(k)) = \theta$  for all  $k$ . For all  $k$  we have

$$\begin{aligned} |x_*(k)|_{X_*} &\geq |x_*(k) - P_{X_*}(x_*(k-1))| - |P_{X_*}(x_*(k-1)) - P_{X_*}(x_*(k))| \\ &= \tan \theta |x_*(k-1)|_{X_*} - |P_{X_*}(x_*(k-1)) - P_{X_*}(x_*(k))| \\ &\geq \tan \theta |x_*(k-1)|_{X_*} - \sup_{\omega_1, \omega_2 \in X_*} |\omega_1 - \omega_2|. \end{aligned} \quad (49)$$

At this point, we need the following conclusion: consider a nonnegative sequence  $\{z_k\}_{k=0}^{\infty}$  with  $z_{k+1} \geq (\tan \theta)z_k - \hat{d}$  and  $\theta > \pi/4$ . Then  $\lim_{k \rightarrow \infty} z_k = \infty$  if  $(\tan \theta - 1)z_0 - \hat{d} > 0$ . Note that  $z_1 - z_0 \geq (\tan \theta - 1)z_0 - \hat{d} > 0$  and then  $z_2 - z_1 \geq (\tan \theta - 1)z_1 - \hat{d} \geq (\tan \theta - 1)z_0 - \hat{d} > 0$ . We can show similarly by induction that  $z_{k+1} - z_k \geq (\tan \theta - 1)z_0 - \hat{d}$  for all  $k$ . The conclusion follows from (49) and the above conclusion.  $\square$

Combining Propositions 5.1, 5.2 and 5.3, we see that  $\pi/4$  is a critical value of the angle error in the approximate projection regarding maintaining bounded iterative states. If  $\theta < \pi/4$ , the system trajectories are uniformly bounded; if  $\theta > \pi/4$ , the trajectories diverge for a special case with one single node and particular approximate projection points; if  $\theta = \pi/4$ , the trajectories of the algorithm with one node are bounded (no longer uniformly with respect to initial conditions) and the property of the trajectories highly depend on the shape of the convex set.

## 6 Numerical Example

In this section, we provide a numerical example which shows that the approximate projected consensus algorithm presented in this paper may lead to a faster convergence than the projected consensus algorithm presented in [27].

Consider a network with three nodes 1, 2 and 3. Suppose  $x_i(k) \in R^2, i = 1, 2, 3$ . Let  $X_1, X_2$  and  $X_3$  be three unit disks with centers  $(1, 0), (-1, 0)$  and  $(0, -1)$ , respectively. Set  $X_i$  corresponds to node  $i$  for  $i = 1, 2, 3$ . The communication graph is completely connected. Take  $a_{11} = a_{22} = a_{33} = 0.5$  and all other weights are 0.25. Here  $\theta_k \equiv 0$ , which is a special case of the approximate projected consensus algorithm.

The three agents have the same initial condition  $x_1(0) = x_2(0) = x_3(0)$ . We consider 2500 initial conditions  $\{(-1.96+0.08p, -1.96+0.08q) | 0 \leq p \leq 49, 0 \leq q \leq 49\}$  equally spaced over the square  $\{(z_1, z_2) | |z_1| \leq 2, |z_2| \leq 2\}$ . The initial conditions from which the approximate projected consensus algorithm with  $\alpha_{i,k} \equiv 0.5$  converges faster than the projected consensus algorithm presented in [27] (compare their distance function  $h(k) = \max_{1 \leq i \leq 3} |x_i(k)|_{X_0}$  at time 2000) are labeled with stars in Figure 4. The decay process of  $h(k)$  of a concrete example with initial condition  $x_1(0) = x_2(0) = x_3(0) = (1.8, 0.8)$  is given in Figure 5 (drawn from time 20), from which we can see that the approximate projected consensus algorithm has better performance from the viewpoint of convergence speed.

## 7 Conclusions

In this paper, we presented an approximate projected consensus algorithm for a network to cooperatively compute the intersection of a serial of convex sets, each of which is known only to a particular node. We allowed each node to only compute an approximate projection which locates in the intersection of the convex cone generated by the current state and all directions with the exact projection direction less than some angle and the half-space containing the current state with its boundary is a supporting hyperplane to its own set at its exact projection point onto its set. Sufficient and/or necessary conditions were obtained for the considered algorithm on how much projection accuracy is required to ensure a global consensus within the intersection set,

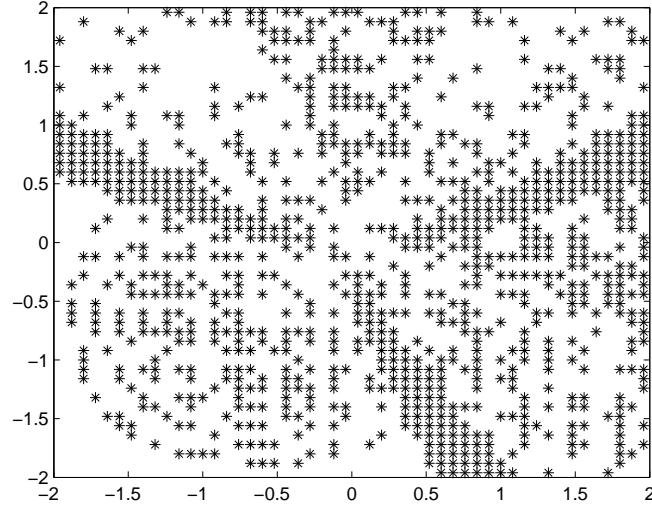


Figure 4: Approximate projected consensus algorithm converges faster than projected consensus algorithm for the initial conditions labeled stars.

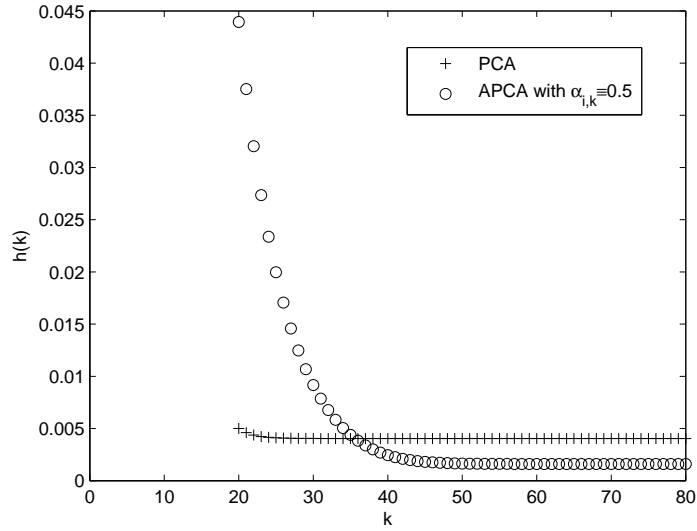


Figure 5: A typical solution for which approximate projected consensus algorithm (APCA) converges faster than projected consensus algorithm (PCA).

under the assumption that the communication graph is uniformly jointly strongly connected. We also showed that  $\pi/4$  is a threshold for the angle error in the projection approximation to ensure a bounded solution for iterative projections. A numerical example was also given indicating that the approximate projected consensus algorithm sometimes achieves better performance than the exact projected consensus algorithm. This implied that, individual optimum seeking may not be so important for optimizing the collective objective.

## References

- [1] R. T. Rockafellar, *Convex Analysis*. New Jersey: Princeton University Press, 1972.
- [2] F. Clarke, Yu.S. Ledyae, R. Stern and P. Wolenski, *Nonsmooth Analysis and Control Theory*. Springer-Verlag, 1998.
- [3] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
- [4] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Princeton Hall, 1989.
- [5] C. Godsil and G. Royle, *Algebraic Graph Theory*. Springer-Verlag, New York, 2001.
- [6] B. T. Polyak, *Introduction to Optimization*. Optimization Software, Inc., New York, 1987.
- [7] J. Tsitsiklis, D. Bertsekas, and M. Athans, “Distributed asynchronous deterministic and stochastic gradient optimization algorithms,” *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803-812, 1986.
- [8] A. Jadbabaie, J. Lin, and A. S. Morse, “Coordination of groups of mobile agents using nearest neighbor rules,” *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988-1001, 2003.
- [9] S. Muthukrishnan, B. Ghosh, and M. Schultz, “First and second order diffusive methods for rapid, coarse, distributed load balancing,” *Theory of Computing Systems*, vol. 31, pp. 331-354, 1998.
- [10] R. Diekmann, A. Frommer, and B. Monien, “Efficient schemes for nearest neighbor load balancing,” *Parallel Computing*, vol. 25, pp. 789-812, 1999.
- [11] G. Shi and Y. Hong, “Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies,” *Automatica*, vol. 45, no. 5, pp. 1165-1175, 2009.
- [12] G. Shi, Y. Hong and K. H. Johansson, “Connectivity and set tracking of multi-agent systems guided by multiple moving leaders,” *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 663-676, 2012.
- [13] M. H. DeGroot, “Reaching a consensus,” *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118-121, 1974.



- [14] D. Acemoglu, A. Ozdaglar and A. ParandehGheibi, "Spread of (mis)information in social networks," *Games and Economic Behavior*, vol. 70, no. 2, pp. 194-227, 2010.
- [15] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655-661, 2005.
- [16] S. Martinez, J. Cortes, and F. Bullo, "Motion coordination with distributed information," *IEEE Control Syst. Mag.*, vol. 27, no. 4, pp. 75-88, 2007.
- [17] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401-420, 2006.
- [18] M. Cao, A. S. Morse and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: a graphical approach," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 575-600, 2008.
- [19] H. G. Tanner, A. Jadbabaie, G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Trans. Autom. Control*, vol. 52, no.5, pp. 863-868, 2007.
- [20] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *IPSN'04*, pp. 20-27, 2004.
- [21] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Stochastic incremental gradient descent for estimation in sensor networks," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, pp. 582-586, 2007.
- [22] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," in *Proc. IEEE Conference on Decision and Control*, Cancun, Mexico, pp. 4185-4190, 2008.
- [23] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. on Decision and Control, Cancun, Mexico*, pp. 4177-4184, 2008.
- [24] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Incremental stochastic subgradient algorithms for convex optimization," *SIAM Journal on Optimization*, vol. 20, no. 2, pp. 691-717, 2009.
- [25] B. Johansson, M. Rabi and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM Journal on Optimization*, vol. 20, no. 3, pp. 1157-1170, 2009.
- [26] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48-61, 2009.
- [27] A. Nedić, A. Ozdaglar and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922-938, 2010.
- [28] A. Nedić and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1757-1780, 2008.

- [29] J. Lu, C. Y. Tang, P. R. Regier, and T. D. Bow, "Gossip algorithms for convex consensus optimization over networks," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2917-2923, 2011.
- [30] J. Lu, P. R. Regier, and C. Y. Tang, "Control of distributed convex optimization," *Proc. IEEE Conference on Decision and Control*, Atlanta, pp. 489-495, 2010.
- [31] G. Shi, K. H. Johansson, and Y. Hong, "Multi-agent systems reaching optimal consensus with directed communication graphs," *Proc. American Control Conference*, San Francisco, pp. 5456-5461, 2011.
- [32] G. Shi and K. H. Johansson, "Randomized optimal consensus of multi-agent systems," <http://arxiv.org/abs/1108.3223>.
- [33] D. Jakovetic, J. Xavier, and J. M. F. Moura, "Cooperative convex optimization in networked systems: augmented lagrangian algorithms with directed gossip communication," <http://arxiv.org/abs/1007.3706>.
- [34] L. Wang and L. Guo, "Robust consensus and soft control of multi-agent systems with noises," *Journal of Systems Science and Complexity*, vol. 21, no.3, pp. 406-415, 2008.
- [35] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, no. 3, pp. 337-404, 1950.
- [36] L. Gubin, B. Polyak, and E. Raik, "The method of projections for finding the common point of convex sets," *U.S.S.R Comput. Math. Math. Phys.*, vol. 7, no. 6, pp. 1211-1228, 1967.
- [37] F. Deutsch, "Rate of convergence of the method of alternating projections," in *Parametric Optimization and Approximation*, B. Brosowski and F. Deutsch, Eds. Basel, Switzerland: Birkhäuser, vol. 76, pp. 96-107, 1983.